MS 31: Rational Krylov Methods and Applications

Part I, 1:45 PM - 3:45 PM	Part II, 4:15 PM - 6:15 PM
Steven Elsworth The Block Rational Arnoldi Algorithm	**Volker Grimm** Rational Krylov Methods in Discrete Inverse Problems
Jörn Zimmerling Rational Krylov Subspaces for Wavefield Applications	**Melina Freitag** Inexact Rational Krylov Methods Applied to Lyapunov Equations
Giampaolo Mele Krylov Methods for Hermitian Nonlinear Eigenvalue Problems	**Davide Palitta** Numerical Methods for Lyapunov Matrix Equations with Banded Symmetric Data
Stefan Güttel Compressing Variable-Coefficient Helmholtz Problems via RKFIT	**Patrick Kürschner** A Comparison of Rational Krylov and Related Low-Rank Methods for Large Riccati Equations

The block rational Arnoldi algorithm

Steven Elsworth   Stefan Güttel

School of Mathematics, The University of Manchester

Sunday 6th May, 2018

Overview

1. Motivation
   - History of block Krylov methods
   - Applications
2. Define a block rational Krylov space
3. The block rational Arnoldi algorithm
   - Continuation vectors
   - Model order reduction example
4. The implicit Q-theorem
5. Rational matrix-valued polynomials
   - RKFUNBs
   - Vector Autoregression

Motivation: History and applications

- Block Lanczos algorithm: Cullum and Donath (1974), Golub and Underwood (1977), Ruhe (1979)
- Block Arnoldi algorithm: Saad (1992)
- Block Krylov subspaces: Gutknecht and Schmelzer (2005, 2009), Frommer, Lund and Szyld (2017)
- Approximate block rational Krylov subspaces: Mach, Pranic and Vandebril (2014)

- Block GMRES: Simoncini and Gallopoulos (1996), Freitag, Kürschner and Pestana (2018)
- Continuous Ricatti equations: Heyouni and Jbilou (2009)
- Model Order Reduction: Abidi, Hached and Jbilou (2014)

Block Krylov spaces

Let $A \in \mathbb{C}^{N \times N}, \mathbf{b} \in \mathbb{C}^{N \times s}$ and suppose the block Krylov matrix $[\mathbf{b}, A\mathbf{b}, \ldots, A^m\mathbf{b}]$ if of full column rank.

- The *classic block Krylov space of order* $m+1$ is defined as

$$\mathcal{K}^{\square}_{m+1}(A, \mathbf{b}) = \left\{ \sum_{k=0}^{m} A^k \mathbf{b} C_k : C_k \in \mathbb{C}^{s \times s} \right\}.$$

- The dimension of this space is $(m+1)s^2$.

Block Krylov spaces

Let $A \in \mathbb{C}^{N\times N}, \mathbf{b} \in \mathbb{C}^{N\times s}$ and suppose the block Krylov matrix $[\mathbf{b}, A\mathbf{b}, \dots, A^m\mathbf{b}]$ if of full column rank.

- The *classic block Krylov space of order* $m+1$ is defined as

$$\mathcal{K}^{\square}_{m+1}(A, \mathbf{b}) = \left\{ \sum_{k=0}^{m} A^k \mathbf{b} C_k : C_k \in \mathbb{C}^{s\times s} \right\}.$$

- The dimension of this space is $(m+1)s^2$.

Definition

Given a nonzero polynomial $q_m \in \mathcal{P}_m$ with roots $\xi_1, \xi_2, \dots, \xi_m \in \overline{\mathbb{C}} \backslash \Lambda(A)$, we define the associated *block rational Krylov space of order* $m+1$ as

$$\mathcal{Q}^{\square}_{m+1}(A, \mathbf{b}, q_m) := q_m(A)^{-1} \mathcal{K}^{\square}_{m+1}(A, \mathbf{b}).$$

Defining block inner-product and block orthogonality [†]

Definition

A mapping $\langle\langle \cdot, \cdot \rangle\rangle : \mathbb{C}^{N\times s} \times \mathbb{C}^{N\times s} \to \mathbb{C}^{s\times s}$ is a *block inner product onto* $\mathbb{C}^{s\times s}$ if $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^{N\times s}$ and $C \in \mathbb{C}^{s\times s}$,

1. $\mathbb{C}^{s\times s}$-linearity: $\langle\langle \mathbf{x}, \mathbf{y}C \rangle\rangle = \langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle C$ and $\langle\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle\rangle = \langle\langle \mathbf{x}, \mathbf{z} \rangle\rangle + \langle\langle \mathbf{y}, \mathbf{z} \rangle\rangle$
2. symmetry: $\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle = \langle\langle \mathbf{y}, \mathbf{x} \rangle\rangle^*$
3. definiteness: $\langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle$ is positive definite if $\mathbf{x}$ has full rank, and $\langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle = O_{s\times s}$ if and only if $\mathbf{x} = O_{N\times s}$.

[†] A. Frommer, K. Lund, and D. Szyld, 2017.

Defining block inner-product and block orthogonality [†]

Definition

A mapping $\langle\langle \cdot, \cdot \rangle\rangle : \mathbb{C}^{N\times s} \times \mathbb{C}^{N\times s} \to \mathbb{C}^{s\times s}$ is a *block inner product onto* $\mathbb{C}^{s\times s}$ if $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^{N\times s}$ and $C \in \mathbb{C}^{s\times s}$,

1. $\mathbb{C}^{s\times s}$-linearity: $\langle\langle \mathbf{x}, \mathbf{y}C \rangle\rangle = \langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle C$ and $\langle\langle \mathbf{x}+\mathbf{y}, \mathbf{z} \rangle\rangle = \langle\langle \mathbf{x}, \mathbf{z} \rangle\rangle + \langle\langle \mathbf{y}, \mathbf{z} \rangle\rangle$
2. symmetry: $\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle = \langle\langle \mathbf{y}, \mathbf{x} \rangle\rangle^*$
3. definiteness: $\langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle$ is positive definite if $\mathbf{x}$ has full rank, and $\langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle = O_{s\times s}$ if and only if $\mathbf{x} = O_{N\times s}$.

- $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\} \subset \mathbb{C}^{N\times s}$ is *block orthonormal* if $\langle\langle \mathbf{x}_i, \mathbf{x}_j \rangle\rangle = \delta_{i,j} I_{s\times s}$, where $\delta_{i,j}$ is the Kronecker delta.
- A mapping $N(\cdot) : \mathbb{C}^{N\times s}(\text{full rank}) \to \mathbb{C}^{s\times s}$ is *scaling quotient* for $\forall \mathbf{x} \in \mathbb{C}^{N\times s}(\text{full rank})$ if $\exists \mathbf{y} \in \mathbb{C}^{N\times s}$ such that

$$\mathbf{x} = \mathbf{y}N(\mathbf{x}) \text{ and } \langle\langle \mathbf{y}, \mathbf{y} \rangle\rangle = I_{s\times s}.$$

[†] A. Frommer, K. Lund, and D. Szyld, 2017.

Constructing a block rational Krylov basis

Input: $A \in \mathbb{C}^{N\times N}$, $\mathbf{b} \in \mathbb{C}^{N\times s}$ of full rank, finite poles $\{\xi_j\}_{j=1}^{m} \subset \mathbb{C} \setminus \Lambda(A)$.

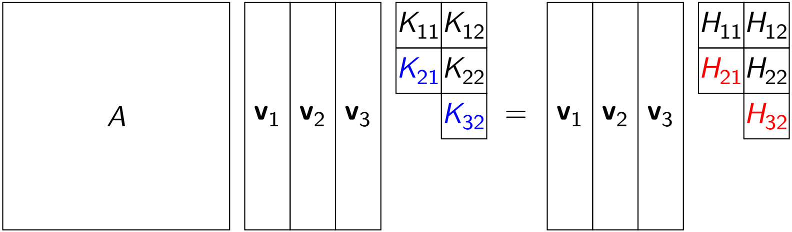
1. $\mathbf{v}_1 := \mathbf{b}N(\mathbf{b})^{-1}$
2. **for** $j = 1, \ldots, m$ **do**
3. Choose a *continuation block vector* $\mathbf{t}_j \in \mathbb{C}^{js\times s}$
4. $\mathbf{w} := (A - \xi_j I)^{-1}\mathbf{V}_j\mathbf{t}_j$
5. **for** $i = 1, \ldots, j$ **do**
6. $C_{i,j} := \langle\langle \mathbf{v}_i, \mathbf{w} \rangle\rangle$
7. Compute $\mathbf{w} := \mathbf{w} - \mathbf{v}_i C_{i,j}$
8. **end for**
9. $C_{j+1,j} := N(\mathbf{w})$
10. $\mathbf{v}_{j+1} := \mathbf{w}C_{j+1,j}^{-1}$
11. Set $\underline{\mathbf{k}_j} := \underline{\mathbf{c}_j} - \underline{\mathbf{t}_j}$ and $\underline{\mathbf{h}_j} := \xi_j\underline{\mathbf{c}_j} - \underline{\mathbf{t}_j}$, where $\underline{\mathbf{t}_j} = [\mathbf{t}_j^T \quad O]^T$.
12. **end for**

Block rational Arnoldi decomposition

Consider a *block rational Arnoldi decomposition* (BRAD)

$$A\mathbf{V}_{m+1}\underline{\mathbf{K}_m} = \mathbf{V}_{m+1}\underline{\mathbf{H}_m}$$

for the space $\mathcal{Q}^{\square}_{m+1}(A, \mathbf{b}, q_m) = q_m(A)^{-1}\mathcal{K}^{\square}_{m+1}(A, \mathbf{b})$. The block upper-Hessenberg matrices $\underline{\mathbf{H}_m}$ and $\underline{\mathbf{K}_m}$ form an unreduced pencil $(\underline{\mathbf{H}_m}, \underline{\mathbf{K}_m})$ with $H_{j+1,j} = \xi_j K_{j+1,j}$.



Unreduced means: At least one of $H_{j+1,j}$ and $K_{j+1,j}$ is nonsingular.

Different choices of continuation block vectors $\mathbf{t}_j$

- `'first'`: $\mathbf{t}_j = \begin{bmatrix} I_{s\times s} & O_{s\times s} & \cdots & O_{s\times s} \end{bmatrix}^T$
- `'last'`: $\mathbf{t}_j = \begin{bmatrix} O_{s\times s} & O_{s\times s} & \cdots & I_{s\times s} \end{bmatrix}^T$
- `'ruhe'`: Given BRAD, subtract $\xi_j \mathbf{V}_{j+1} \underline{\mathbf{K}_j}$ from both sides,

$$(A - \xi_j I)\mathbf{V}_{j+1}\underline{\mathbf{K}_j} = \mathbf{V}_{j+1}(\underline{\mathbf{H}_j} - \xi_j \underline{\mathbf{K}_j}).$$

Compute $(\underline{\mathbf{H}_j} - \xi_j \underline{\mathbf{K}_j}) = Q\underline{R}$, then

$$\mathbf{V}_{j+1} Q \quad Q^* \underline{\mathbf{K}_j} R^{-1} = (A - \xi_j I)^{-1} \mathbf{V}_{j+1} Q \begin{bmatrix} I_{js\times js} \\ O_{s\times js} \end{bmatrix}.$$

Define

$\mathbf{t}_1 = I_{s\times s}$ and $\mathbf{t}_j = Q($`:, end-s+1:end`$)$.

Different choices of continuation block vectors $\mathbf{t}_j$

- `'first'`: $\mathbf{t}_j = \begin{bmatrix} I_{s\times s} & O_{s\times s} & \cdots & O_{s\times s} \end{bmatrix}^T$
- `'last'`: $\mathbf{t}_j = \begin{bmatrix} O_{s\times s} & O_{s\times s} & \cdots & I_{s\times s} \end{bmatrix}^T$
- `'ruhe'`: Given BRAD, subtract $\xi_j \mathbf{V}_{j+1} \underline{\mathbf{K}_j}$ from both sides,

$$(A - \xi_j I)\mathbf{V}_{j+1}\underline{\mathbf{K}_j} = \mathbf{V}_{j+1}(\underline{\mathbf{H}_j} - \xi_j \underline{\mathbf{K}_j}).$$

  Compute $(\underline{\mathbf{H}_j} - \xi_j \underline{\mathbf{K}_j}) = Q\underline{R}$, then

$$\mathbf{V}_{j+1}Q \;\; Q^* \underline{\mathbf{K}_j} R^{-1} = (A - \xi_j I)^{-1} \mathbf{V}_{j+1} Q \begin{bmatrix} I_{js\times js} \\ O_{s\times js} \end{bmatrix}.$$

  Define

$$\mathbf{t}_1 = I_{s\times s} \text{ and } \mathbf{t}_j = Q(:,\ \texttt{end-s+1:end}).$$

- Can show: with repeated poles, `'ruhe'` = `'last'`

Choice of continuation vector matters!

Example Problem: Approximate transfer function

$$H(s) = c^*(sE - A)^{-1}\mathbf{b}$$

over frequency range $i[0, 40]$, for nonsymmetric matrices $\{A, E\} \subset \mathbb{R}^{N\times N}$ and block vector $\mathbf{b} \in \mathbb{R}^{N\times 2}$, where $N = 11730$[†].

Method: Compute orthonormal block rational Krylov basis $\mathbf{V}_m$. Project $A_m = \mathbf{V}_m^* A \mathbf{V}_m$ and $E_m = \mathbf{V}_m^* E \mathbf{V}_m$, and define approximation

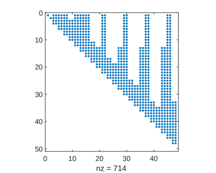
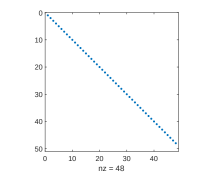
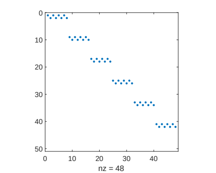
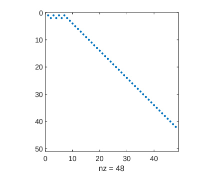
$$H_m(s) = (c^*\mathbf{V}_m)(sE_m - A_m)^{-1}(\mathbf{V}_m^*\mathbf{b}).$$

[†] G. Lassaux and K. Willcox, *Model reduction for active control design using multiple-point Arnoldi methods*, AIAA Paper, 616 (2003), pp. 1–11.

Choice of continuation vector matters!

Cyclically repeat 4 equispaced poles on $i[0, 40]$, until dimension $\mathrm{m} = 24$.
Use CGS without reorthogonalisation.

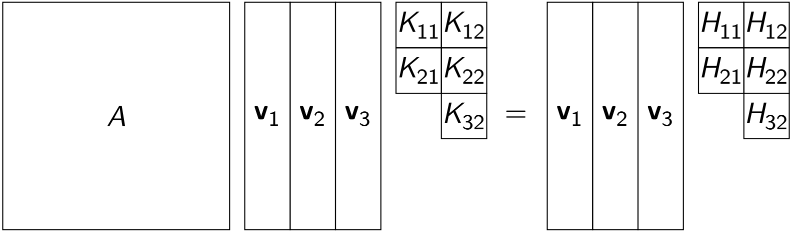
- cond: $\kappa(\mathbf{X})$, condition number of block basis before orthogonalisation.
- orth: $\|\mathbf{V}^T\mathbf{V} - I\|_2$, orthogonality of computed block basis.

'ruhe' $\mathrm{p} = 1$	'last' $\mathrm{p} = 1$	'last' $\mathrm{p} = 4$	'almost-last' $\mathrm{p} = 4$
			
cond: **4.2×10^{3}**	cond: 8.7×10^{5}	cond: 8.7×10^{5}	cond: 6.0×10^{4}
orth: **2.6×10^{-11}**	orth: 1.1×10^{-8}	orth: 7.1×10^{-10}	orth: 1.8×10^{-10}

Essentially equal BRADs

Definition

Two orthonormal BRADs, $A\mathbf{V}_{m+1}\underline{\mathbf{K}_m} = \mathbf{V}_{m+1}\underline{\mathbf{H}_m}$ and $A\widehat{\mathbf{V}}_{m+1}\underline{\widehat{\mathbf{K}}_m} = \widehat{\mathbf{V}}_{m+1}\underline{\widehat{\mathbf{H}}_m}$, are **essentially equal** if there exists a unitary block diagonal matrix $\mathbf{D}_{m+1} \in \mathbb{C}^{(m+1)s\times(m+1)s}$, and a block upper-triangular nonsingular matrix $\mathbf{T}_m \in \mathbb{C}^{ms\times ms}$, such that $\widehat{\mathbf{V}}_{m+1} = \mathbf{V}_{m+1}\mathbf{D}_{m+1}$, $\underline{\widehat{\mathbf{H}}_m} = \mathbf{D}^*_{m+1}\underline{\mathbf{H}_m}\mathbf{T}_m$, and $\underline{\widehat{\mathbf{K}}_m} = \mathbf{D}^*_{m+1}\underline{\mathbf{K}_m}\mathbf{T}_m$.



Implicit Q-theorem allows rerunning

Theorem ([Mach et al., 2014][E. & Güttel, 2018])

Consider an orthonormal BRAD, $A\mathbf{V}_{m+1}\underline{\mathbf{K}_m} = \mathbf{V}_{m+1}\underline{\mathbf{H}_m}$ *with poles* $\{\xi_j\}_{j=1}^m \subset \overline{\mathbb{C}} \setminus \Lambda(A)$.
The block-orthonormal matrix $\mathbf{V}_{m+1}$ *and the pencil* $(\underline{\mathbf{H}_m}, \underline{\mathbf{K}_m})$ *are essentially uniquely determined by* $\mathbf{v}_1$ *and the poles* $\xi_1, \dots, \xi_m$.

Consider a BRAD $A\mathbf{V}_{m+1}\underline{\mathbf{K}_m} = \mathbf{V}_{m+1}\underline{\mathbf{H}_m}$.
Given $\widetilde{A} \in \mathbb{C}^{\widetilde{N}\times\widetilde{N}}$ and $\widetilde{\mathbf{v}}_1 \in \mathbb{C}^{\widetilde{N}\times s}$, we can construct

$$\widetilde{\mathbf{V}}_{m+1} = [\widetilde{\mathbf{v}}_1, \cdots, \widetilde{\mathbf{v}}_{m+1}]$$

such that

$$\widetilde{A}\widetilde{\mathbf{V}}_{m+1}\underline{\mathbf{K}_m} = \widetilde{\mathbf{V}}_{m+1}\underline{\mathbf{H}_m}.$$

Rational matrix-valued polynomials

- We can show that

$$\mathbf{v}_{j+1} = R_j(A) \circ \mathbf{v}_1 \text{ for } j = 1, 2, \ldots, m,$$

where $R_j(z) = q_m(z)^{-1}(C_{j,0} + zC_{j,1} + \cdots + z^m C_{j,m})$, and $C_{j,i} \in \mathbb{C}^{s \times s}$ are encoded in $(\underline{\mathbf{H}_j}, \underline{\mathbf{K}_j})$. †

- RKFUNB is a representation of a rational matrix-valued function of the form $(\underline{\mathbf{H}_m}, \underline{\mathbf{K}_m}, \texttt{coeffs})$ where `coeffs` is an array of square $s \times s$ matrices. The rational function is defined as

$$R(z) = R_0(z)\texttt{coeffs}(1) + \cdots + R_m(z)\texttt{coeffs}(m+1).$$

† If $P(z) = C_0 + zC_1 + \cdots + z^m C_m$, where $\{C_0, \ldots, C_m\} \subset \mathbb{C}^{s \times s}$.
Let $A \in \mathbb{C}^{N \times N}, \mathbf{b} \in \mathbb{C}^{N \times s}$, then $P(A) \circ \mathbf{b} = C_0 + A\mathbf{b}C_1 + \cdots + A^m \mathbf{b} C_m$.

What is Vector Autoregression?

A stationary mean centred multivariate time series **y** can be modelled by $VAR(p)$ process if

$$\mathbf{y}_t = \mathbf{y}_{t-1} C_1 + \cdots + \mathbf{y}_{t-p} C_p + \varepsilon_t,$$

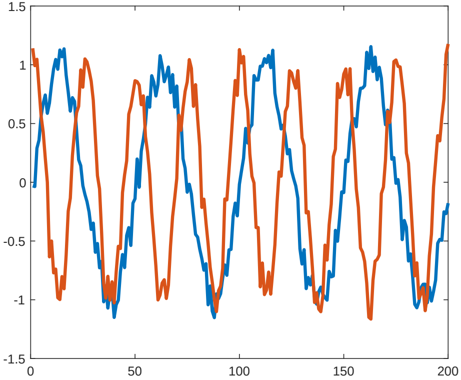
where $C_1, \ldots, C_p \in \mathbb{C}^{s \times s}$ and ε is multivariate white noise.

What is Vector Autoregression?

A stationary mean centred multivariate time series $\mathbf{y}$ can be modelled by $VAR(p)$ process if

$$\mathbf{y}_t = \mathbf{y}_{t-1} C_1 + \cdots + \mathbf{y}_{t-p} C_p + \varepsilon_t,$$

where $C_1, \ldots, C_p \in \mathbb{C}^{s \times s}$ and ε is multivariate white noise.



Vector Autoregression as an RKFUNB

Define

$$A = \begin{pmatrix} \mathbf{0} & I_{N-1\times N-1} \\ 0 & \mathbf{0}^T \end{pmatrix}, \ \mathbf{b} = \mathbf{y}_t, \ \mathtt{xi} = [\underbrace{\infty, \dots, \infty}_{p-1}] \text{ and } F = A^p.$$

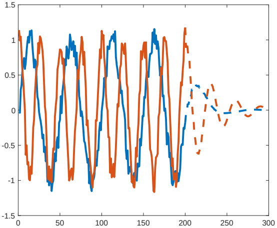
Using block bilinear form

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle = \mathbf{x}^* \begin{pmatrix} I_{N-p\times N-p} & O_{N-p\times p} \\ O_{p\times N-p} & O_{p\times p} \end{pmatrix} \mathbf{y},$$

we minimise

$$\|A^p\mathbf{b} - (\mathbf{v}_1 C_1 + \cdots + \mathbf{v}_p C_p)\|^2 = \|F\mathbf{b} - r(A) \circ \mathbf{b}\|^2.$$

```
r = rkfitb(F, A, b, xi, param);
pred = mu + r(A, b);
```



Conclusions and Future Work

- 'ruhe' continuation vector is a good default continuation strategy.
- Implicit Q-Theorem allows rerunning a decomposition.
- RKFUNBs provide a representation of rational matrix-valued polynomials.

- S. Elsworth, S. Güttel, *The block rational Arnoldi algorithm*, In preparation.
- MATLAB Rational Krylov Toolbox: `http://rktoolbox.org`.
- Block generalisation of RKFIT.
- ARMA and VARMA models.

EPSRC Engineering and Physical Sciences Research Council

sabisu®

siam®